

N O T I C E

THIS DOCUMENT HAS BEEN REPRODUCED FROM
MICROFICHE. ALTHOUGH IT IS RECOGNIZED THAT
CERTAIN PORTIONS ARE ILLEGIBLE, IT IS BEING RELEASED
IN THE INTEREST OF MAKING AVAILABLE AS MUCH
INFORMATION AS POSSIBLE

AgRISTARS

"Made available under NASA sponsorship
in the interest of early and wide dis-
semination of Earth Resource Survey
Program information and without liability
for any use made thereof."

Foreign Commodity Production Forecasting

FC-L1-04073
JSC-17129

NASA-CR-161057

**A Joint Program for
Agriculture and
Resources Inventory
Surveys Through
Aerospace
Remote Sensing**

June 1981

U.S. CORN AND SOYBEANS EXPLORATORY EXPERIMENT SUMMARY REPORT

E82-10088

CR-161057

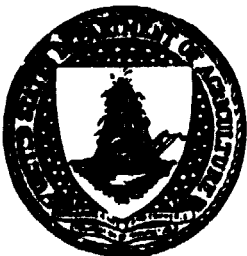
J. G. Carnes

(E82-10088) US CORN AND SOYBEANS
EXPLORATORY EXPERIMENT (Lockheed Engineering
and Management) 29 p HC A03/MF A01 CSCL 02C

N82-21640

Unclas
G3/43 00088

Lockheed Engineering and Management Services Company, Inc.
1830 NASA Road 1, Houston, Texas 77058



NASA



**Lyndon B. Johnson Space Center
Houston, Texas 77058**

1. Report No. JSC-17129; FC-L1-04073		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle U.S. Corn and Soybeans Experiment — Summary Report				5. Report Date June 1981	
				6. Performing Organization Code	
7. Author(s) J. G. Carnes				8. Performing Organization Report No. LEMSCO-16315	
9. Performing Organization Name and Address Lockheed Engineering and Management Services Company, Inc. 1830 NASA Road 1 Houston, Texas 77058				10. Work Unit No.	
				11. Contract or Grant No.	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Lyndon B. Johnson Space Center Houston, Texas 77058 Technical Monitor: J. D. Erickson				13. Type of Report and Period Covered NAS 9-15800	
				14. Sponsoring Agency Code	
15. Supplementary Notes					
16. Abstract This report summarizes the results from the U.S. Corn/Soybeans Exploratory Experiment which was completed during FY 1980. The experiment consisted of two parts: the Classification Procedures Verification Test and the Simulated Aggregation Test. Results are presented for evaluations of labeling, proportion estimation, and aggregation procedures.					
17. Key Words (Suggested by Author(s)) aggregation FCPF AgRISTARS proportion estimation analyst labeling remote sensing corn/soybeans				18. Distribution Statement	
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 28	
22. Price*					

U.S. CORN AND SOYBEANS EXPLORATORY EXPERIMENT
SUMMARY REPORT

Job Order 72-415

This report describes the U.S. Corn/Soybean Exploratory Experiment of the Foreign Commodity Production Forecasting project of the AgRISTARS program.


PREPARED BY

J. G. Carnes

APPROVED BY

NASA

LOCKHEED


J. D. Erickson, Manager
Foreign Commodity Production
Forecasting Project


B. L. Carroll, Manager
Crop Applications Department

LOCKHEED ENGINEERING AND MANAGEMENT SERVICES COMPANY, INC.

Under Contract NAS 9-15800

For

Earth Resources Applications Division
Space and Life Sciences Directorate
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

June 1981

PRECEDING PAGE BLANK NOT FILMED

PREFACE

The Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing is an 8-year program of research, development, evaluation, and application of aerospace remote sensing for agricultural resources, which began in fiscal year 1980. This program is a cooperative effort of the National Aeronautics and Space Administration, the U.S. Agency for International Development, and the U.S. Departments of Agriculture, Commerce, and the Interior.

The work which is the subject of this document was performed within the Earth Resources Applications Division, Space and Life Sciences Directorate, at the Lyndon B. Johnson Space Center, National Aeronautics and Space Administration. Under Contract NAS 9-15800, personnel of Lockheed Engineering and Management Services Company, Inc., performed the tasks which contributed to the completion of this research.

EXECUTIVE SUMMARY

In order to perform the 1981 U.S. Corn/Soybeans Pilot Experiment, a segment-level proportion estimation procedure and a procedure for aggregating segment-level proportion estimates to produce large-area acreage estimates were needed.

The U.S. Corn/Soybeans Exploratory Experiment described in this report was performed to obtain a better understanding of the performance of the technology before proceeding to the pilot experiment evaluation under a larger set of agricultural conditions. The experiment investigated the use of techniques for estimating corn and soybeans proportions which were similar to the techniques used for small grains proportion estimation. An evaluation was made of an aggregation procedure which was developed for further evaluation in the pilot experiment under different conditions.

The results of the proportion estimation procedure evaluation indicated that, while the labeling part of the procedure worked quite well, the machine classification part was not effective in improving the proportion estimates based on labeling information alone. The aggregation procedure performed quite well in producing large-area estimates of acreage and production.

The evaluation indicated that the technology was capable of producing generally acceptable area estimates for corn and soybeans. There were problems with certain aspects of the technology, but the problems were understood and could be addressed in on-going development activities. The development of the proportion estimation procedure for the pilot experiment relied to a great extent on the understanding of the technology resulting from the exploratory experiment. The evaluation of the aggregation procedure showed that it was ready for use in the pilot experiment.

PRECEDING PAGE BLANK NOT FILMED

CONTENTS

Section	Page
1. PURPOSE AND SCOPE.....	1-1
2. INTRODUCTION.....	2-1
2.1 <u>AgRISTARS PROGRAM</u>	2-1
2.2 <u>FCPF PROJECT</u>	2-1
2.3 <u>U.S. CORN/SOYBEANS EXPLORATORY EXPERIMENT</u>	2-1
3. EXPERIMENT DESCRIPTION.....	3-1
3.1 <u>CLASSIFICATION PROCEDURE VERIFICATION TEST DESCRIPTION</u>	3-1
3.2 <u>SIMULATED AGGREGATION - TEST DESCRIPTION</u>	3-5
4. EXPERIMENT RESULTS.....	4-1
4.1 <u>CLASSIFICATION PROCEDURES VERIFICATION TEST</u>	4-1
4.2 <u>SIMULATED AGGREGATION TEST RESULTS</u>	4-4
5. CONCLUSIONS AND RECOMMENDATIONS.....	5-1
6. REFERENCES.....	6-1

TABLES

Table		Page
1.	SUMMARY OF RESULTS FOR THE CLASSIFICATION PROCEDURES VERIFICATION TEST.....	4-2
2.	COMPARISON OF LABELING ACCURACY FOR THE CLASSIFICATION PROCEDURES VERIFICATION TEST AND THE SIMULATED AGGREGATION TEST.....	4-7

FIGURES

Figure		
1.	Map showing locations of the segments used in the Classification Procedure Verification Test.....	3-2
2.	Diagram showing procedure for processing segments for the Classification Procedures Verification Test.....	3-3
3.	Diagram showing the major steps in the labeling procedure for the Classification Procedures Verification Test.....	3-4
4.	Map showing locations of the segments used in the Simulated Aggregation Test.....	3-7
5.	Proportion estimation errors as a function of the true proportion for both corn and soybeans.....	4-3
6.	Proportion errors when only pure pixels are used to determine the proportions.....	4-5
7.	Summary of results for the Simulated Aggregation Test.....	4-6
8.	Coefficients of variation resulting from reduced acquisition rates for acreage and production for three states.....	4-9

1. PURPOSE AND SCOPE

The purpose of this report is to summarize the results from the U.S. Corn/Soybeans Exploratory Experiment. This experiment was completed during fiscal year (FY) 1980. It was conducted as part of the Foreign Commodity Production Forecasting (FCPF) project of the Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing (AgRISTARS) program.

2. INTRODUCTION

2.1 AgRISTARS PROGRAM

The AgRISTARS program is an 8-year program of research, development, and evaluation of the application of aerospace remote sensing in monitoring agricultural resources. The program began in FY 1980. The AgRISTARS program is a cooperative effort of the National Aeronautics and Space Administration (NASA), the U.S. Agency for International Development (AID), and the U.S. Departments of Agriculture, Commerce, and the Interior (USDA, USDC, and USDI). The goal of this program is to determine the usefulness, cost, and extent to which aerospace remote sensing data can be used by the USDA to improve the objectivity, reliability, and timeliness of information required to carry out USDA missions (ref. 1).

2.2 FCPF PROJECT

An important component of the AgRISTARS program is the FCPF project. The objective of the FCPF project is to develop and test procedures for using aerospace remote sensing technology to provide more objective, timely, and reliable crop production forecasting in foreign areas. To develop technology for use in foreign areas, the FCPF project builds upon existing remote sensing technology and extends this technology to additional crops and regions (ref. 2).

During the first year of the FCPF project, two exploratory experiments were performed using U.S. data to develop and evaluate techniques. These experiments were the U.S./Canada Wheat and Barley Exploratory Experiment (1) and the U.S. Corn/Soybean Exploratory Experiment. This report presents the results from the U.S. Corn/Soybean Exploratory Experiment.

2.3 U.S. CORN/SOYBEAN EXPLORATORY EXPERIMENT

The overall objective for the U.S. Corn/Soybean Exploratory Experiment was to evaluate segment-level labeling and proportion estimation techniques and to test a procedure for making aggregated estimates of area and production. These two components plus a yield estimation component form the nucleus of a system

for making large area crop acreage and production estimates. The results of these evaluations provide the basis for establishing baseline procedures for subsequent experiments and for adaptation of the procedures for use in foreign regions.

The U.S. Corn/Soybean Exploratory Experiment was the first attempt to adapt the segment-level proportion estimation techniques developed for small grains to other crops. The segment-level proportion estimates were obtained by labeling selected pixels from the segment as training for a maximum likelihood classifier. The results from the classification were corrected for bias using an independent set of labeled pixels. The pixel labeling was done using an objective procedure based on labeling techniques developed during previous experiments. This marks the first time an objective procedure was used to label pixels instead of relying on the experience and insight of highly trained analysts to obtain pixel labels.

In order to test the weighted aggregation procedure proposed for the FCPF project, a simulation test was performed as part of this exploratory experiment. The weighted aggregation procedure was set up to make the best use of historical data to stabilize large area aggregated estimates in regions where there is a high rate of data loss or where there are large classification errors.

3. EXPERIMENT DESCRIPTION

Two tests were performed as part of this exploratory experiment. The first test was the Classification Procedures Verification Test (CPVT). The second test was the Simulated Aggregation Test (SAT).

3.1 CLASSIFICATION PROCEDURES VERIFICATION TEST DESCRIPTION

The CPVT consisted of labeling and proportion estimation on 25 segments from four agrophysical units (APU's) in the U.S. Corn Belt using Landsat data from the 1978 crop year. The locations of the segments used in the CPVT are shown in figure 1.

The two objectives of this test were (1) to determine the accuracy of the newly developed objective labeling procedure and recommend improvements for use in the SAT and (2) to determine the accuracy of the proportion estimation procedure.

The procedure used to process the segments for this test is shown in figure 2. Using Landsat and ancillary data, an objective labeling procedure (ref. 3) was used to label two sets of pixels from each segment. The major steps in the labeling procedure are shown in figure 3. The procedure is set up to provide increasingly detailed labeling information at each step in the procedure. The first step consists of a decision tree labeling logic which is used to separate the pixels into cropland and noncropland. The pixels labeled cropland in the first step are separated into summer crops and "other crops" in the second step. This step also uses a decision tree labeling logic. In the third step, a greenness/brightness scatter plot for the separation acquisition is used for separating the summer crop pixels into corn and soybeans.

The first set of analyst-labeled pixels (Type I dots) is used as training for a clustering algorithm which grouped all of the pixels in the segment into clusters on the basis of their spectral values. Each of the resulting clusters is labeled as corn, soybeans, or "other" using the labeled Type I dot closest to the mean of the cluster. On the basis of the means and variances for each

U.S. CORN BELT PROCEDURES VERIFICATION TEST

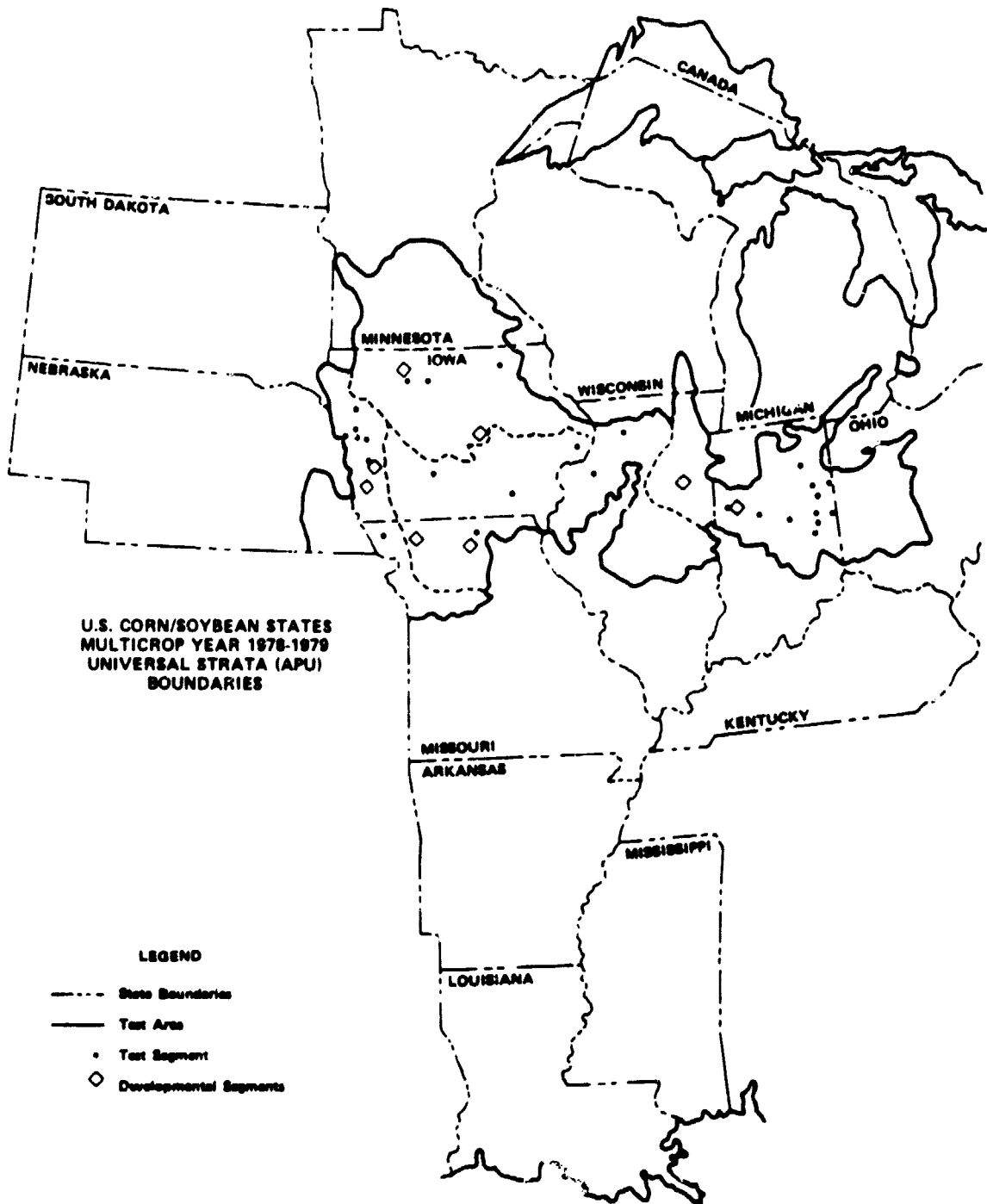


Figure 1.- Map showing locations of the segments used in the Classification Procedures Verification Test.

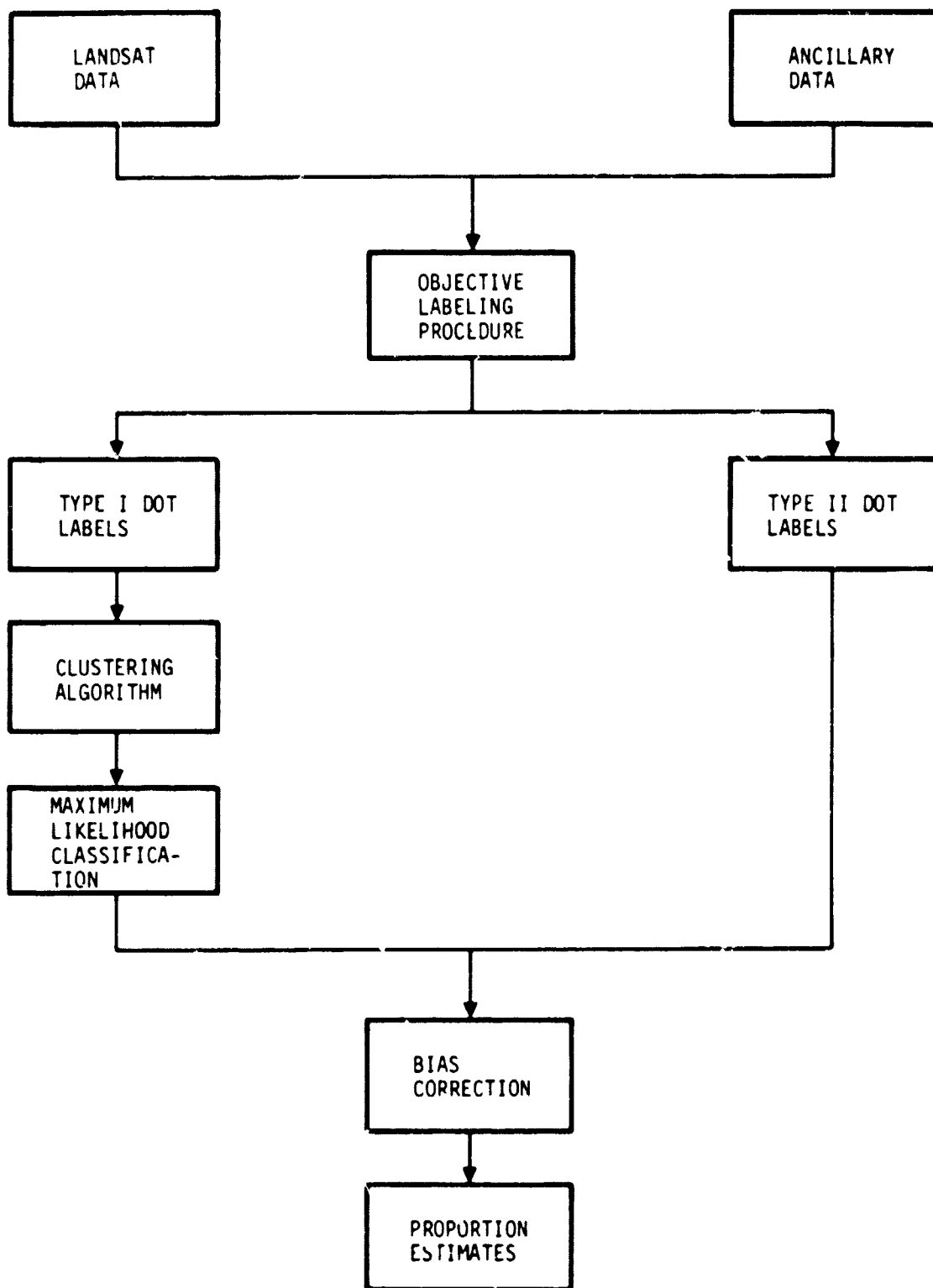


Figure 2.- Diagram showing procedure for processing segment for the Classification Procedures Verification test.

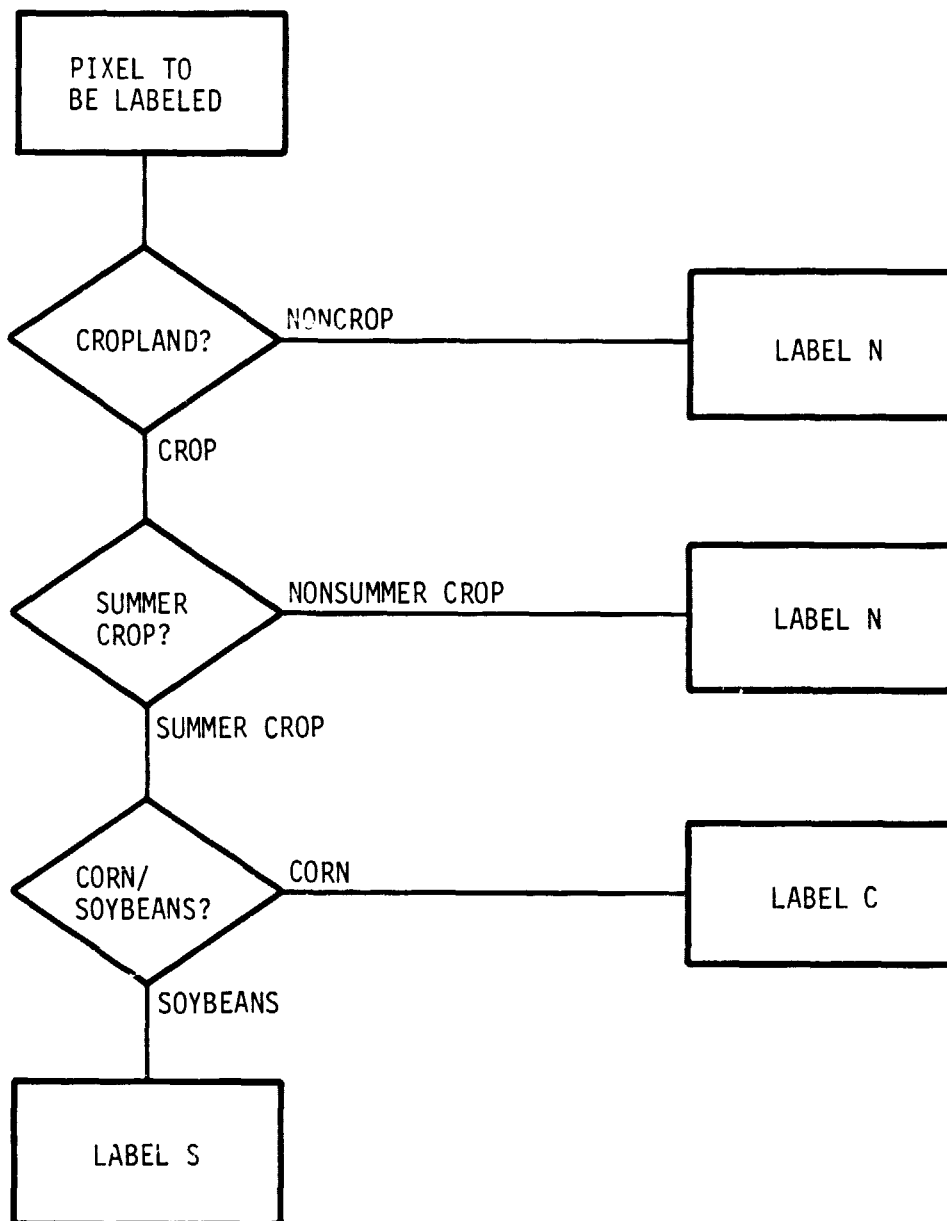


Figure 3.- Diagram showing the major steps in the labeling procedure for the Classification Procedures Verification test.

cluster, a maximum likelihood classification of every pixel in the segment is performed. Using the second set of analyst-labeled dots (Type 2 dots) as a random sample of the segment, the proportion based on the classification is corrected for any bias introduced by the classification process.

The segments in the CPVT were processed independently by three groups of analysts. Each segment was processed by at least two of the groups. The test followed a rigid experiment design so that analysis of variance techniques could be used to determine if the quality of the labeling and proportion estimation results were dependent on the group doing the labeling or on the APU in which the segment was located (ref. 4). All of the evaluations were performed by comparing the labeling and classification results to the digitized ground-truth crop inventories.

An error characterization study was performed to determine if any changes could be made in the labeling procedure to improve the accuracy of the labeling. These changes were incorporated into the procedure before it was used in the second test.

An evaluation was made to determine the effectiveness of the classification procedure in producing proportion estimates.

3.2 SIMULATED AGGREGATION TEST DESCRIPTION

The second test performed as part of this exploratory experiment was the SAT. It involved labeling and proportion estimation for 88 segments in the Corn Belt using Landsat data from the 1978 crop year. These results were used to estimate the variability of proportion estimates for a simulation test of the aggregation procedure. There were two parts to the test. The first part was the actual labeling and classification for the segments. The objective of this part was to evaluate the labeling and proportion estimation procedures as they were modified following the CPVT and to provide estimates of the variability introduced by the labeling and proportion estimation procedures for the aggregation procedure simulation test. The second part of the test was the actual aggregation procedure simulation itself. The objectives of the simulation were to verify that

the sample allocation procedure provided correct sample allocations among the strata, to validate the aggregation and variance estimation logic, and to determine the effect of random nonresponse on aggregated estimates produced by the procedure.

The labeling procedure used in the SAT was essentially the same as that used in the CPVT. The changes made as a result of the CPVT were mainly improvements in the clarity of the procedure. The proportion estimation procedure was modified from the procedure used in the CPVT. Based on a study performed by Supporting Research (ref. 5) and on resource considerations, the decision was made not to perform the bias correction on the initial proportion estimates in the SAT. Therefore, the proportion estimation procedure involved labeling of the Type I dots, classification of the segment, and proportion estimation by enumeration of pixels in the class of interest.

The aggregation procedure tested in the second part of the SAT consisted of a technique for using historical data to compensate for the loss of data in a particular stratum (ref. 6). The technique involves a weighting procedure which places more reliance on historical data as the classification results become less reliable because of data loss or errors in the classification results.

The 88 segments in the SAT were each processed once. Twenty-three of the segments had been processed in the CPVT. These were processed in the SAT, but by a different analyst group. Thirty-five additional segments with ground-truth inventories were processed and used in the evaluations. The locations of the segments used in this test are shown in figure 4. Evaluation of the labeling and proportion estimation accuracies were performed using the segments for which ground-truth information was available.

The simulation test of the aggregation procedure was performed by setting up an allocation of 204 simulated segments in 12 strata in the states of Illinois, Indiana, and Iowa. Historical data were used to determine the mean crop proportions for each stratum. The distribution of segment-level crop proportions

SIMULATED AGGREGATION TEST

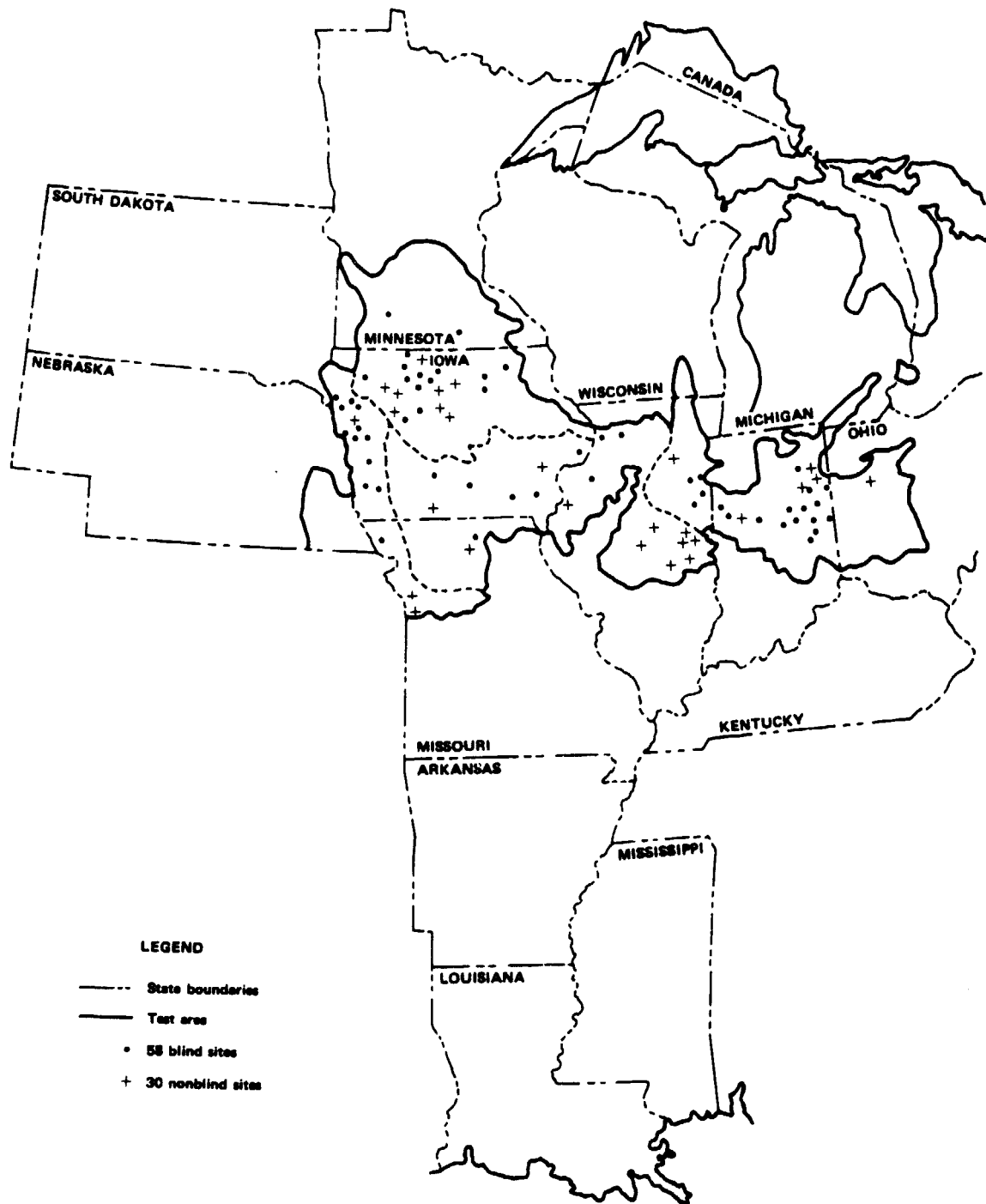


Figure 4.- Map showing locations of the segments used in the Simulated Aggregation Test.

7

within strata was determined from the historical variability and from the empirical variances observed in the classification results. State-level historical data were used to determine mean yields, and the distribution of yield estimates was determined using ESS yield-model variance. A Monte Carlo simulation was performed in which segments were randomly designated as "lost". For each loss rate, 100 simulations were performed to obtain aggregated estimates of production.

4. EXPERIMENT RESULTS

4.1 CLASSIFICATION PROCEDURES VERIFICATION TEST RESULTS

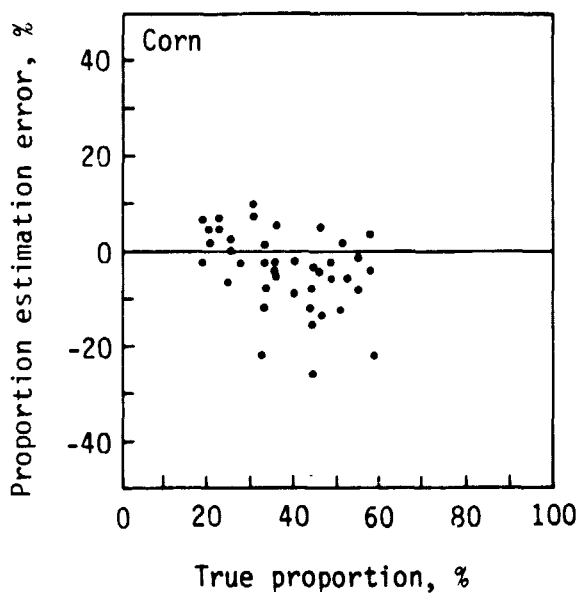
In the CPVT, statistical tests were performed to determine if there was a significant difference in the quality of the labeling and proportion estimation results due to the group performing the processing or the region in which the segment was located (ref. 7). The measures of quality used were dot labeling accuracy, percentage of correct classification, and proportion estimation error. A regional difference was observed for the dot labeling accuracy for soybeans. The labeling of soybeans was significantly worse in a predominantly corn-producing region than in the regions where soybeans were more prevalent. A group effect was found in the dot labeling accuracy for corn. One group produced significantly less accurate dot labeling for corn. Investigation showed that the difference was due to a difference in the way the group placed the separation line on the scatter plots for corn and soybeans.

The labeling accuracies for the CPVT are shown in table 1. The labeling accuracy is comparable to the small grains labeling accuracies previously achieved during the LACIE. The labeling for Type I dots was better than for Type II dots. This difference results from the fact that the Type I dots are required to be spectrally pure, while the Type II dots can be spectrally mixed. It is, therefore, natural to expect better labeling accuracy on dots which are representative of a particular crop, rather than a mixture of signatures from more than one crop.

The proportion estimation errors as a function of the true proportion are shown for both corn and soybeans in figure 5. The average proportion of corn in the segments was 38 percent. The machine processing procedure underestimated the corn proportion by an average of 4 percent. The average proportion of soybeans was 28 percent. The procedure underestimated the soybeans proportion by 6 percent. All of the bias and half of the variability in the proportion estimation errors were the result of dot labeling errors. The proportion estimates produced by the procedure were not any better than estimates obtained by using the Type II dots as a random sample. Therefore, the machine processing (i.e., clustering and classification) did not improve the results.

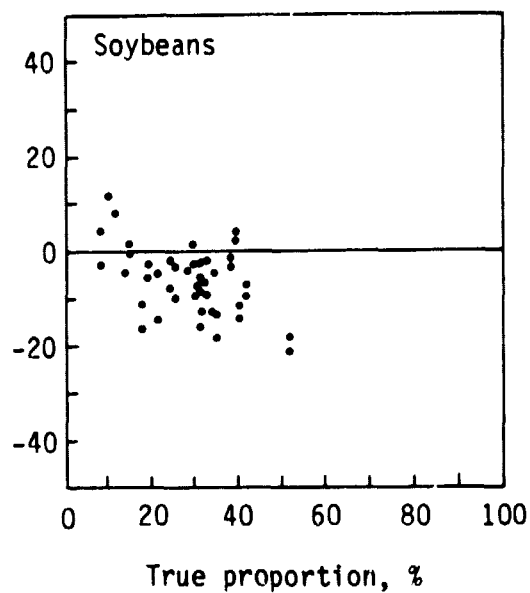
TABLE 1.- SUMMARY OF DOT LABELING RESULTS FOR THE
CLASSIFICATION PROCEDURES VERIFICATION TEST

Ground-truth category	Percent correctly labeled	
	Type 1 dots	Type 2 dots
Corn	83	73
Soybeans	79	64
Other	93	86
All categories	86	75



Mean error = -4%
 Standard deviation = 8%
 RMSE = 9%

(a) Corn — bias-corrected estimate.



Mean error = -6%
 Standard deviation = 7%
 RMSE = 9%

(b) Soybeans — bias-corrected estimate.

Figure 5.- Proportion estimation errors as a function of the true proportion for both corn and soybeans.

Since the labeling and classification accuracies were much better for spectrally pure pixels than for mixed pixels, a study was made on the segments in this test to determine if accurate proportion estimates could be obtained from classification information for spectrally pure pixels. In order to perform the study, each of the pure pixels was given its ground-truth label, and a proportion estimate was made using only these pixels. Figure 6 shows the proportion estimation errors for two criteria for pixel purity. Pixels which meet the "one-half pixel" purity criterion are at least one-half pixel from the field boundaries. Pixels which meet the "one pixel" criterion are at least one pixel from the field boundaries. The results indicate that proportion estimates based only on pure pixels can be biased and have a great deal of variability. In the data set used in this test, the corn estimates showed a positive bias.

4.2 SIMULATED AGGREGATION TEST RESULTS

In the SAT, the labeling accuracy improved over the accuracy in the CPVT (ref. 8). Table 2 shows a comparison of the labeling accuracies in the two tests. The improvement in the labeling accuracy for the second test was due to changes in the labeling procedure recommended on the basis of the first test and to an improved procedure for selecting acquisitions.

The proportion estimation results for the SAT are shown in figure 7. The results for soybeans proportion estimation were comparable to those obtained in the CPVT. The average soybeans proportion in the segments was 30 percent. The procedure underestimated the soybeans proportions by an average of 8 percent. For corn, the average proportion was 41 percent. In the SAT, the procedure overestimated the corn proportions by 5 percent, while in the CPVT, the proportions were underestimated by 4 percent. The change in bias between the two tests is due to the fact that a bias correction was not performed in the SAT. The proportions based on the pixel-level classification results are too high because the classification procedure was trained using only spectrally pure pixels. Training with only pure pixels produces a classification which is representative of the pure areas of the segment, rather than of the entire segment. As the pure pixel studies showed, this will produce a positive bias in the classification results.

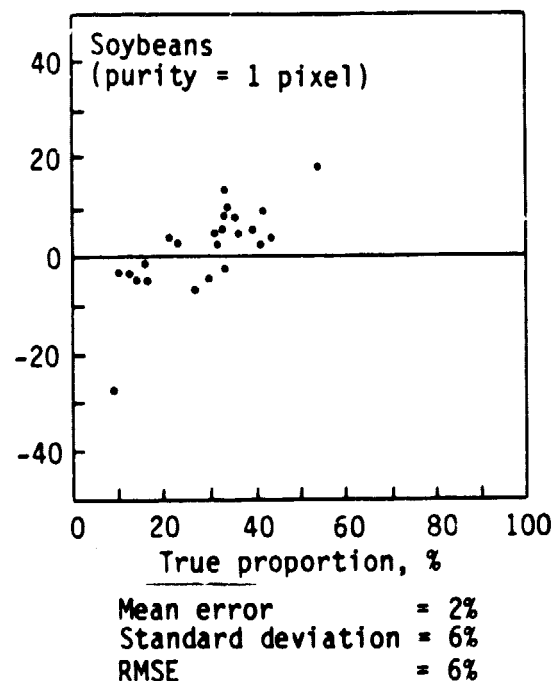
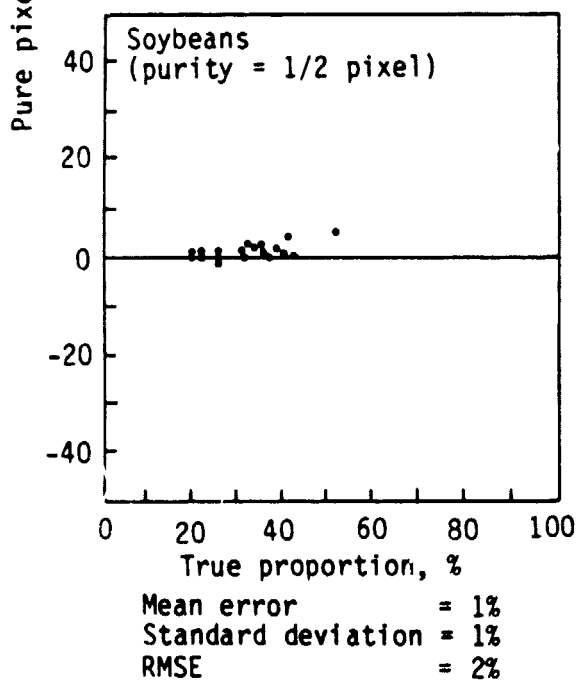
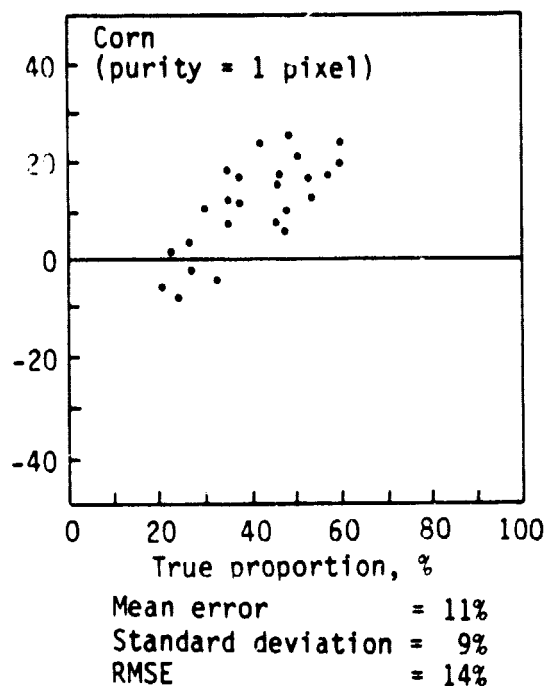
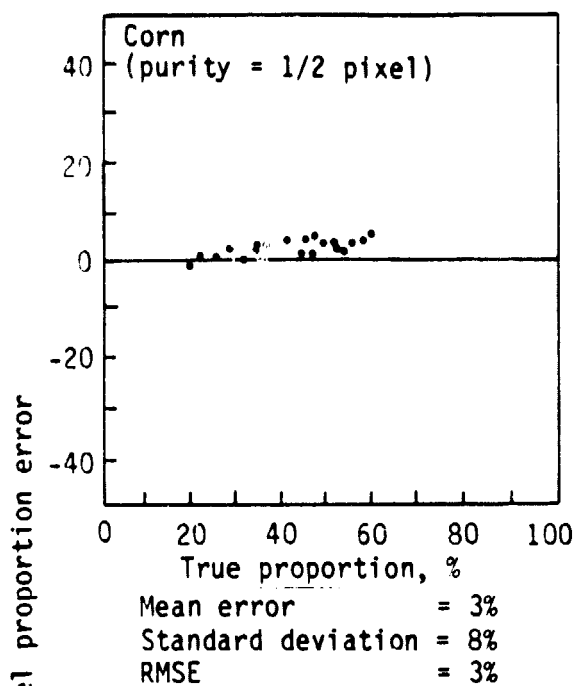
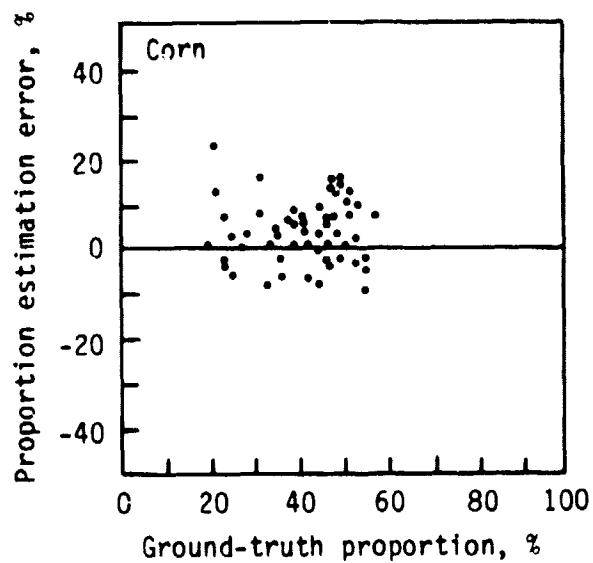


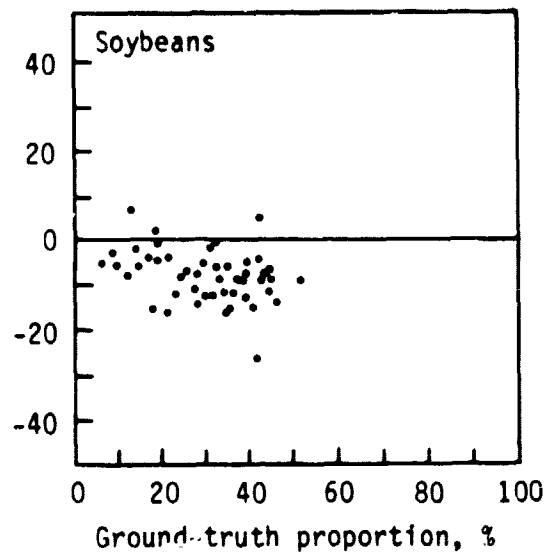
Figure 6.- Proportion errors when only pure pixels are used to determine the proportions.

TABLE 2.- COMPARISON OF LABELING ACCURACY FOR
CPVT AND SAT TESTS

Ground-truth categories	Percent correctly labeled	
	CPVT (Type I dots)	SAT
Corn	86	93
Soybeans	79	88
Other	93	96
All categories	86	92



Mean error = 5%
 Standard deviation = 7%
 RMSE = 8%



Mean error = -8%
 Standard deviation = 6%
 RMSE = 9%

Figure 7.- Summary of results for the simulated Aggregation Test.

The simulation tests of the sampling and aggregation procedures (ref. 9) were set up to provide large area production estimates with a coefficient of variation (CV) of 5 percent for both corn and soybeans at an acquisition rate of 100 percent. The aggregation procedure was tested to determine if the CV estimates computed by the procedure were correct, if nonresponse introduced any bias into the aggregated estimates, and if the CV's at reduced response rates were reasonable.

The simulation tests showed that the allocation procedure was producing estimates with CV's in good agreement with the expected value of 5 percent (CV = 4.7% for corn and CV = 5.2% for soybeans). The tests of the weighted aggregation procedure demonstrated that the procedure introduced no bias into the aggregated area and production estimates for acquisition rates as low as 10 percent. Figure 8 shows the CV's resulting from reduced acquisition rates for area and for production. These variances are reasonable, and the average CV estimates produced by the procedure correspond closely to the CV's of the simulated sample.

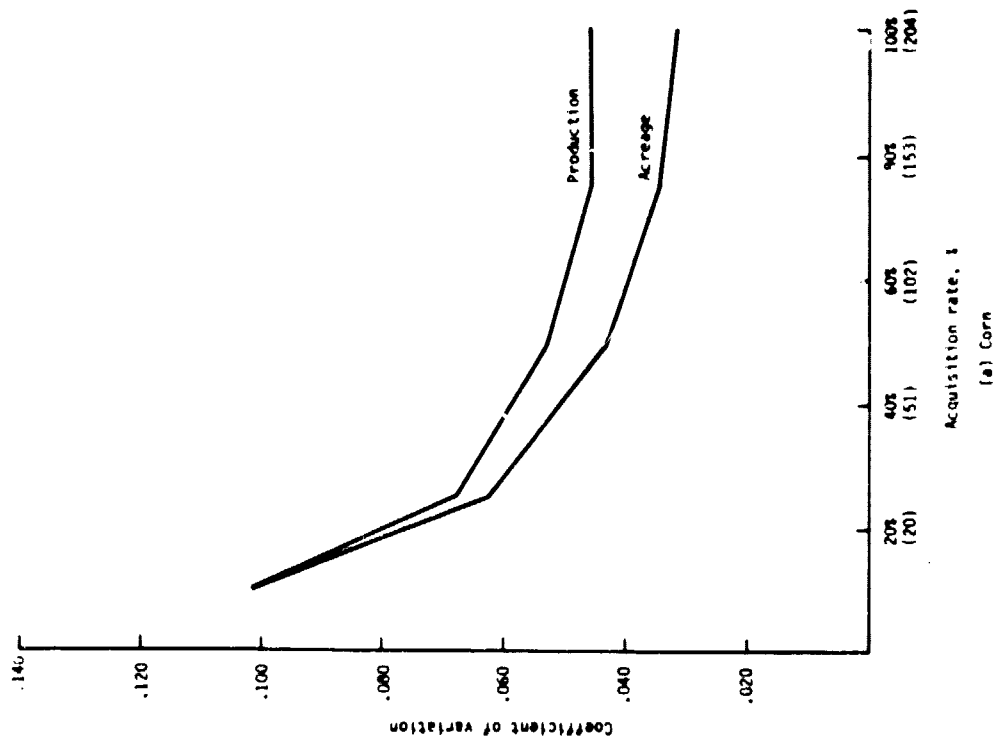
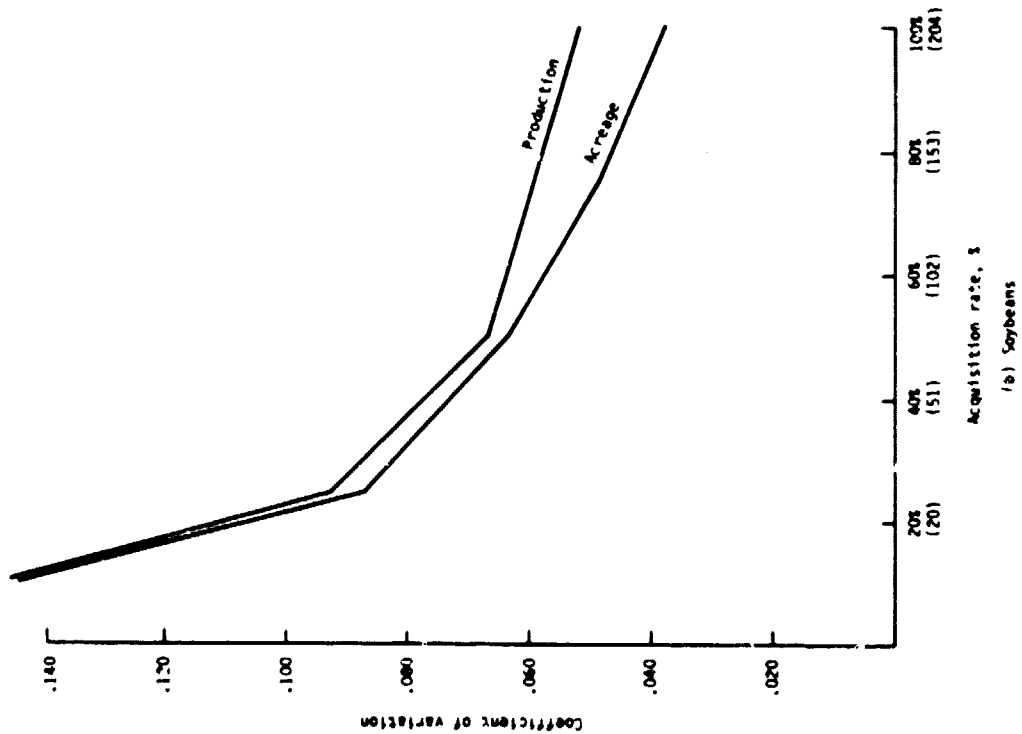


Figure 8.- Coefficients of variation resulting from reduced acquisition rates for acreage and production for three states.

5. CONCLUSIONS AND RECOMMENDATIONS

The results from the labeling evaluations indicate that the corn/soybeans labeling procedure performs very well in the U.S. Corn Belt with full-season (after tasseling) Landsat data. The procedure should be readily adaptable to corn/soybeans labeling required for subsequent exploratory experiments or pilot tests.

The machine classification procedures evaluated in this experiment were not effective in improving the proportion estimates. The corn proportions produced by the machine procedures had a large bias when the bias correction was not performed. This bias was caused by the manner in which the machine procedures handled spectrally impure pixels. Alternatives to the machine processing techniques used in this experiment should be investigated to see if more effective techniques can be found.

The simulation test indicated that the weighted aggregation procedure performed quite well. Although further work can be done to improve both the simulation tests and the aggregation procedure, the results of this test show that the procedure should serve as a useful baseline procedure in future exploratory experiments and pilot tests.

6. REFERENCES

1. Technical Program Plan - AgRISTARS. January 15, 1980.
2. Foreign Commodity Production Forecasting Project Implementation Plan for Fiscal Years 1981 and 1982. FC-JO-C0614, JSC-16828, January 9, 1981.
3. Dailey, C. L.; and Abotteen, K. M.: Corn/Soybeans Decision Logic Development and Testing. LEMSCO-14811 (JSC-16380), October 1980.
4. LACIE Transition Project Accuracy Assessment Fiscal Year 79 Interim Plan, LEC-13753 (JSC-13770), August 1979.
5. Havens, K. A.; and Abotteen, K. M.: Evaluation of Three-Category Classification. LEC-13498 (JSC-16012), August 1979.
6. Feiveson, A. H.: Weighted Ratio Estimation for AgRISTARS. Technical Memorandum (to be published).
7. Carnes, J. G.; and Baird, J. E.: Evaluation of Results of U.S. Corn and Soybeans Exploratory Experiment - Classification Procedures Verification Test. LEMSCO-14386 (JSC-16339), September 1980.
8. Davidson, S. A.: Segment-Level Evaluation of the Simulated Aggregation Test: U.S. Corn and Soybean Exploratory Experiment. LEMSCO-15116 (JSC-16820), September 1980.
9. U.S. Corn/Soybeans Exploratory Experiment Final Report, April, 1981 (to be published).